

1/22

## FIG. 1

BLASTP 2.2.1 [Jul-12-2001]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= genscan2

{416 letters}

Database: ncbi-nr

897,014 sequences; 280,886,335 total letters

Searching.....done

Sequences producing significant alignments: (bits) Value

pir  JC1512 biliary glycoprotein H - mouse	81	2e-14
pir  JC1506 biliary glycoprotein B - mouse	81	2e-14
pir  A39037 carcinoembryonic antigen m <sub>m</sub> CGM2 precursor - mouse >g...	79	9e-14
ref NP_036056.1  (NM_011926) CEA-related cell adhesion molecule ...	79	9e-14
pir  JC1509 biliary glycoprotein E - mouse	73	5e-12
ref NP_001758.1  (NM_001767) CD2 antigen (p50), sheep red blood ...	73	6e-12
ref NP_113943.1  (NM_031755) carcinoembryonic antigen-related ce...	72	8e-12
pir  RWHUC2 T-cell surface glycoprotein CD2 precursor - human >g...	72	8e-12
gb AAAS1946.1  (M16336) CD2 surface antigen (Homo sapiens)	72	8e-12
ref NP_291021.1  (NM_033543) hypothetical protein R29124_1 [Homo...	72	1e-11
pir  JC1507 biliary glycoprotein C - mouse	71	2e-11
emb CAA47697.1  (X67280) biliary glycoprotein [Mus musculus]	71	2e-11
pir  S34338 biliary glycoprotein F - mouse >gi 312586 emb CAA476...	71	2e-11
pir  JC1511 biliary glycoprotein G - mouse	71	2e-11

FIG. 1 (contd.)

3/22  
FIG. 2

>pir||JC1512 biliary glycoprotein H - mouse

Length = 341

Score = 80.9 bits (198), Expect = 2e-14

Identities = 54/168 (32%), Positives = 86/168 (51%), Gaps = 9/168 (5%)

Query: 73 RDKPVTVVQSIGTEVIGTLR---PDYRDRIRLFENGSLLLSDLQLADEGTYEVEISITD 128

+ PV+ I +V GT + P + R ++ NGSLL+ + + D G Y +E+ TD

Subjct: 69 KGNPVSTNAEIVHQVTGTVKTTTGAHSRETVYSNGSLLIQRVTVKDTGVYTIEM--TD 126

Query: 129 DTFTG-EKTINLTVDVPISRQPQLVASTTVLELSEAFTLNC SHENGTKPSYTWLKDGPL 187

+ F E T+ V P+++P + V +TTV EL ++ TL C N + WL + + L

Subjct: 127 ENFRRTEATVQFHVHQPVTPQSLQVTNTTVKEL-DSVLTCL-SNDIGANIQWLFNSQL 184

Query: 188 LNDSRMLLSPDQKVLTITRVLIMEDDDLYSCMVNPISQGRSLPVKITV 235

RM LS + +L I + ED Y C + NP+S RS +K+ +

Subjct: 185 QLTERMTLSQNNSLRIDPIKREDAGEYQCEISNPVSVKRSNSIKLDI 232

4/22

## FIG. 3

Score	E
Sequences producing significant alignments:	(bits) Value
gi 483306 pir  JC1506 biliary glycoprotein B - mouse	79 1e-13
gi 111207 pir  A39037 carcinoembryonic antigen mCGM2 precu...	77 3e-13
gi 483312 pir  JC1512 biliary glycoprotein H - mouse	77 4e-13
gi 13937381 ref NP_036056.1  (NM_011926) CEA-related cell a...	75 1e-12
gi 228710 prf  1809184A pregnancy-specific glycoprotein (Ra...	70 5e-11
gi 483307 pir  JC1507 biliary glycoprotein C - mouse	70 6e-11
gi 16117775 ref NP_291021.1  (NM_033543) hypothetical prote...	69 8e-11
gi 483309 pir  JC1509 biliary glycoprotein E - mouse	69 9e-11
gi 312582 emb CAA47695.1  (X67278) biliary glycoprotein (Mu...	69 1e-10
gi 483311 pir  JC1511 biliary glycoprotein G - mouse	68 2e-10

5/22  
FIG. 4

>gi|483306|pir||JC1506 biliary glycoprotein B - mouse

Length = 278

Score = 78.6 bits (192), Expect = 1e-13

Identities = 54/168 (32%), Positives = 86/168 (51%), Gaps = 9/168 (5%)

Query: 73 RDKPVTVVQSIGTEVIGTLR---PDYRDRIRLFENGSLLSQLQLADEGTYEVEISITD 128

+ PV+ I +V GT + P + R ++ NGSLL+ + + D G Y +E +TD

Sbjct: 69 KGNPVSTNAEIVHQVTGYNKTTGPAHSGRETVYSNGSLLIQRVTVKDTGVYTIE--MTD 126

Query: 129 DTF-TGEKTINLTVDVPISRQPQLVASTTVLELSEAFTLNCSHENGTKPSYTWLKDGPL 187

+ F E T+ V P++P + V +TTV EL ++ TL C N + WL + + L

Sbjct: 127 ENFRRTEATVQFHVHQPVTPQPSLQVTNTTVKEL-DSVTLTCL-SNDIGANIQWLSQL 184

Query: 188 LNDSRMLLSPDQKVLTITRVLMEDDDLYSCVVENPISQVRSLPVKITV 235

RM LS + +L I + ED Y C + NP+S RS +K+ +

Sbjct: 185 QLTERMTLSQNNISLRIDPIKREDAGEYQCEISNPVSVKRSNSIKLDI 232

6/22  
FIG. 5

1 ATGAAGAGAG AAAGGGGAGC CCTGTCCAGA GCCTCCAGGG CCCTGCGCCT TGCTCCTTTT  
m k r e r g a l s r a s r a l r l a p f

61 GTCTACCTTC TTCTGATCCA GACAGACCC CTGGAGGGGG TGAACATCAC CAGCCCCGTG  
v y l l l i q t d p l e g v n i t s p v

121 CGCCTGATCC ATGGCACCGT GGGGAAGTCG GCTCTGCTTT CTGTGCAGTA CAGCAGTACC  
r l i h g t v g k s a l l s v q y s s t

181 AGCAGCGACA GGCGCTGTAGT GAAGTGGCAG CTGAAGCGGG ACAAGCCAGT GACCGTGGTG  
s s d r p v v k w q l k r d k p v t v v

241 CAGTCCATTG GCACAGAGGT CATCGGCACC CTGCGGCCTG ACTATCGAGA CCGTATCCGA  
q s i g t e v i g t l r p d y r d r i r

301 CTCTTGAAA ATGGCTCCCT GCTTCTCAGC GACCTGCAGC TGGCCGATGA GGGCACCTAT  
l f e n g s l l l s d l q l a d e g t y

361 GAGGTCGAGA TCTCCATCAC CGACGACACC TTCACTGGGG AGAAGACCAT CAACCTTACT  
e v e i s i t d d t f t g e k t i n l t

421 GTAGATGTGC CCATTCGAG GCCACAGGTG TTGGTGGCTT CAACCACTGT GCTGGAGCTC  
v d v p i s r p q v l v a s t t v l e l

481 AGCGAGGCCT TCACCTTGAA CTGCTCACAT GAGAATGGCA CCAAGCCCAG CTACACCTGG  
sea f t l n c s h e n g t k p s y t w

541 CTGAAGGATG GCAAGCCCCT CCTCAATGAC TCGAGAATGC TCCTGTCCCC CGACCAAAAG  
i k d g k p l l n d s r m l l s p d q k

601 GTGCTCACCA TCACCCGCGT GCTCATGGAG GATGACGACC TGTACAGCTG CATGGTGGAG  
v l t i t r v l m e d d d l y s c m v e

661 AACCCCATCA GCCAGGGCCG CAGCCTGCCT GTCAAGATCA CCGTATACAG AAGAAGCTCC  
n p i s q g r s l p v k i t v y r r s s

721 CTTTACATCA TCTTGTCTAC AGGAGGCATC TTCCCTCTTG TGACCTTGGT GACAGTCTGT  
l y i i l s t g g i f l l v t l v t v c

781 GCCTGCTGGA AACCTCCAA AAGGAAACAG AAGAAGCTAG AAAAGCAAAA CTCCCTGGAA  
a c w k p s k r k q k k l e k q n s l e

841 TACATGGATC AGAATGATGA CCGCCTGAAA CCAGAAGCAG ACACCCCTCCC TCGAAGTGGT  
       y m d q n d d r l k p e a d t l p r s g

901 GAGCAGGAAC GGAAGAACCC CATGGCACTC TATATCCTGA AGGACAAGGA CTCCCCGGAG  
       e q e r k n p m a l y i l k d k d s p e

961 ACCGAGGAGA ACCCGGCCCGG AGAGCCTCGA AGCGCGACGG AGCCCGGCCCGC GCCCGGCTAC  
       t e e n p a p e p r s a t e p g p p g y

1021 TCCGTGTCTC CCGCCGTGCC CGGCCGCTCG CCAGGGCTGC CCATCCGCTC TGCCCGCCGC  
       s v s p a v p g r s p g l p i r s a r r

1081 TACCCGCGCT CCCCAGCGCG CTCCCCAGCC ACCGGCCGGA CACACTCGTC GCCGCCAGG  
       y p r s p a r s p a t g r t h s s p p r

1141 GCCCCGAGCT CGCCCCGGCCG CTGGCGCAGC GCCTCGCCGA CACTGGGGAC TGCGGGCTG  
       a p s s p g r s r s a s r t l r t a g v

1201 CACATAATCC GCGAGCAAGA CGAGGCCGGC CCGTGGAGA TCAGCCCTG AGCCGCCCTG  
       h i i r e q d e a g p v e i s a

1261 GGATCCCCTG AGAGGCCGCC CGGGCTCGC GCCAGTGGCC CGGGGGAAAG CTGGGGCTGG  
 1321 GAAGCCCGGG CGCGCGCGC TGGGACGAG GGGAGGTCCC GGGGGGGCGC TGGTGTCTCG  
 1381 GGTGTGAACG TGTATGAGCA TGCAGACG GAGGCGGGTG CGCGGAGGCG GCAGTGTGA  
 1441 TATGGTAAAC CCGGGTCGCA TTTGCTTCG GTTACTGGC TGTGTCTCA CTTGGTATAG  
 1501 GTTGTGCCCT CTTAGGACCA CATAGATTAT TACATTCTG GCCCAATACC CAAAAGGGTT  
 1561 TTATGGAAAC TAACATCACT AACCTAACCC CCGTGAATAT CCTGTGCTCT TCCTAGGGAG  
 1621 CTGTGTTTT TCCCACCCAC CACCCCTCCC TCTGAACAA TGCCCTGAGTG CTGGGGCACT  
 1681 TTTTTTTTTT TTTTTTTTTT GCAAGTTCAAG ATTAGAGAGG CCACTTTCCC  
 1741 AGAATCCACA GCTGCACTAA GCTAAGGAGA AGCCAGATGC CGGTTACTGG GTGTGCAGGG  
 1801 GCTGTTCTGA GCTGGGGGAA TCATTGTGAA GGCCTCTTC CCTGGGCACC TGGTACCTGG  
 1861 GGACCTACAA GGTGGTGAGG GAAGGGTACG AGTACATTCC TTTTCCCTCT GACCTGGGCG  
 1921 CTAGCAAGGG CAAAGAACCC GAGCCTGCCA GCTTGGCCTC CTCCACAGC CTCCCTCGGA  
 1981 GGCATGCCAT GCCAAGCACT CTTTCTGTCT CTGTTCATGA ATAAA

FIG. 5(contd.)



# FIG. 6(contd.)

9/22

INSP052-exon3F

gi 127121 CGACACCTTCACTGGGAGAAGACCATCAACCTTACTGTAGATGgtaaa. 127164

|||||||||||||||||||||||||>>>>

384 CGACACCTTCACTGGGAGAAGACCATCAACCTTACTGTAGATG..... 427

INSP052-exon3F cont

INSP052-exon2R cont

gi 127164 ....ctcagTGCCCATTTCGAGGCCACAGGTGTTGGTGGCTTCAACCACT 127922

717 >>>|||||||||||||||||||||||||

427 .....TGCCCATTTCGAGGCCACAGGTGTTGGTGGCTTCAACCACT 468

INSP052-exon2R

gi 127923 GTGCTGGAGCTCAGCGAGGCCTTCACCTTGAAC TGCTCACATGAGAATGG 127972

|||||||||||||||||||||||||

469 GTGCTGGAGCTCAGCGAGGCCTTCACCTTGAAC TGCTCACATGAGAATGG 518

gi 127973 CACCAAGCCCAGCTACACCTGGCTGAAGGATGGCAAGGCCCTCCTCAATG 128022

|||||||||||||||||||||

519 CACCAAGCCCAGCTACACCTGGCTGAAGGATGGCAAGGCCCTCCTCAATG 568

gi 128023 ACTCGAGAATGCTCCTGTCCCCCGACCAAAAGGTGCTCACCATCACCCGC 128072

|||||||||||||||||||||

569 ACTCGAGAATGCTCCTGTCCCCCGACCAAAAGGTGCTCACCATCACCCGC 618

gi 128073 GTGCTCATGGAGGATGACGACCTGTACAGCTGCATGGTGGAGAACCCCAT 128122

|||||||||||||||||||||

619 GTGCTCATGGAGGATGACGACCTGTACAGCTGCATGGTGGAGAACCCCAT 668

gi 128123 CAGCCAGGGCCGCAGCCTGCCTGTCAAGATCACCGTATACAGtgag... 128163

|||||||||||||||||>>> 295

669 CAGCCAGGGCCGCAGCCTGCCTGTCAAGATCACCGTATACA..... 709

INSP052-exon3R cont

gi 128163 .cctagGAAGAAGCTCCCTTACATCATCTTGTCTACAGGAGGCATCTC 128502

>>>|||||||||||||||||

709 .....GAAGAAGCTCCCTTACATCATCTTGTCTACAGGAGGCATCTC 753

INSP052-exon3R

gi 128503 CTCTTGTGACCTTGGTGACAGTCTGTGCCTGCTGGAAACCTCCAAAAG 128552

|||||||||||||||||

754 CTCTTGTGACCTTGGTGACAGTCTGTGCCTGCTGGAAACCTCCAAAAG 803

**FIG. 6(contd.)****10/22**

gi 128552 gtctg.....cacagGAAACAGAAGAAGCTAGAAAAGCAAAACTCCCTGG 129108

>>>> 521 >>>>||||||||||||||||||||||||||||||||||||||||

803 .....GAAACAGAAGAAGCTAGAAAAGCAAAACTCCCTGG 838

gi 129109 AATACATGGATCAGAATGATGACCGCCTGAAACCAGAAGGttag.....t 129147

||||||||||||||||||||||||||||||||||||||||||>>>> 286 >

839 AATACATGGATCAGAATGATGACCGCCTGAAACCAGAAG..... 877

gi 129147 gcagCAGACACCCTCCCTCGAAGTGGTGAGCAGGAACGGAAGAACCCAT 129479

>>>>||||||||||||||||||||||||||||||||||||||||||||

877 ....CAGACACCCTCCCTCGAAGTGGTGAGCAGGAACGGAAGAACCCAT 923

gi 129480 GGCACACTATATCCTGAAGGACAAGGttag.....tgcaGACTCCCCGG 130461

||||||||||||||||||>>>> 947 >>>>||||||||||

924 GGCACACTATATCCTGAAGGACAAG.....GACTCCCCGG 958

gi 130462 AGACCGAGGAGAACCCGGCCCCGGAGCCTCGAAGCGCGACGGAGCCCGGC 130511

||||||||||||||||||||||||||||||||||||||||||||||||

959 AGACCGAGGAGAACCCGGCCCCGGAGCCTCGAAGCGCGACGGAGCCCGGC 1008

gi 130512 CCGCCCGGCTACTCCGTGTCTCCCGCCGTGCCCGGCCGCTGCCGGGGCT 130561

||||||||||||||||||||||||||||||||||||||||||||||||

1009 CCGCCCGGCTACTCCGTGTCTCCCGCCGTGCCCGGCCGCTGCCGGGGCT 1058

gi 130562 GCCCATCCGCTCTGCCCGCCGCTACCGCGCTCCCCAGCGCGCTCCCCAG 130611

||||||||||||||||||||||||||||||||||||||||||||||

1059 GCCCATCCGCTCTGCCCGCCGCTACCGCGCTCCCCAGCGCGCTCCCCAG 1108

gi 130612 CCACCGGGCGGACACACTCGTCGCCGCCAGGGCCCCGAGCTGCCCGGC 130661

||||||||||||||||||||||||||||||||||||||||||||||

1109 CCACCGGGCGGACACACTCGTCGCCGCCAGGGCCCCGAGCTGCCCGGC 1158

gi 130662 CGCTCGCGCAGCGCCTCGCGCACACTGCGGACTGCCGGCGTGCACATAAT 130711

||||||||||||||||||||||||||||||||||||||||||

1159 CGCTCGCGCAGCGCCTCGCGCACACTGCGGACTGCCGGCGTGCACATAAT 1208

gi 130712 CCGCGAGCAAGACGAGGCCGGCCGGTGGAGATCAGCGCCTGA 130754

||||||||||||||||||||||||||||||||||||||

1209 CCGCGAGCAAGACGAGGCCGGCCGGTGGAGATCAGCGCCTGA 1251

11/22  
**FIG. 7**

1 ACAAGTTGT ACAAAAAAGC AGGCTTCGCC ACCATGAAGA GAGAAAGGGG AGCCCTGTCC  
m k r e r g a l s

61 AGAGCCTCCA GGGCCCTGCG CCTTGCTCCT TTTGCTTACCC TTCTTCTGAT CCAGACAGAC  
r a s r a l r l a p f v y l l l i q t d

121 CCCCTGGAGG GGGTGAACAT CACCAGCCCC GTGCGCCTGA TCCATGGCAC CGTGGGGAAAG  
p l e g v n i t s p v r l i h g t v g k

181 TCGGCTCTGC TTTCTGTGCA GTACAGCAGT ACCAGCAGCG ACAGGCCTGT AGTGAAGTGG  
s a l l s v q y s s t s s d r p v v k w

241 CAGCTGAAGC GGGACAAGCC AGTGACCGTG GTGCAGTCCA TTGGCACAGA GGTCACTCGGC  
q l k r d k p v t v v q s i g t e v i g

301 ACCCTGCGGC CTGACTATCG AGACCGTATC CGACTCTTG AAAATGGCTC CCTGCTTCTC  
t l r p d y r d r i r l f e n g s l l l

361 AGCGACCTGC AGCTGGCCGA TGAGGGCACC TATGAGGTG AGATCTCCAT CACCGACGAC  
s d l q l a d e g t y e v e i s i t d d

421 ACCTTCACTG GGGAGAAGAC CATCACCTT ACTGTAGATG TGCCCATTC GAGGCCACAG  
t f t g e k t i n l t v d v p i s r p q

481 GTGTTGGTGG CTTCAACCAC TGTGCTGGAG CTCAGCGAGG CCTTCACCTT GAACTGCTCA  
v l v a s t t v i e l s e a f t l n c s

541 CATGAGAATG GCACCAAGCC CAGCTACACC TGGCTGAAGG ATGGCAAGCC CCTCCTCAAT  
h e n g t k p s y t w l k d g k p l l n

601 GACTCGAGAA TGCTCCTGTC CCCCGACCAA AAGGTGCTCA CCATCACCCG CGTGCTCATG  
d s r m l l s p d q k v l t i t r v l m

661 GAGGATGACG ACCTGTACAG CTGCATGGTG GAGAACCCCA TCAGCCAGGG CCCGAGCCTG  
. e d d d l y s c m v e n p i s q g r s l

721 CCTGTCAAGA TCACCGTATA CAGAAGAAGC TCCCACCATC ACCATCACCA TTGAAACCCA  
p v k i t v y r r s s h h h h h h -

781 GCTTTCTTGT ACAAAAGTGGT

12/22

## FIG. 8

Molecule: pENTR-INSPO52- EC-6HIS, 3005 bps DNA Circular  
File Name: pENTR-INSPO52-6HIS.cm5, dated 21 Feb 2003

Description: Ligation of Cons-6His.SEQ into pENTR-attL1-attL2

### Molecule Features:

Type	Start	End	Name	Description
MARKER	21			pENTR-F1 primer
MARKER	110		C attL1	
GENE	136	873	INSPO52-EC-6HIS	
MARKER	888		attL2	
MARKER	1001	C		pENTR-R1 primer
GENE	1100	1909	KanR	
REGION	2030	2669	ori	

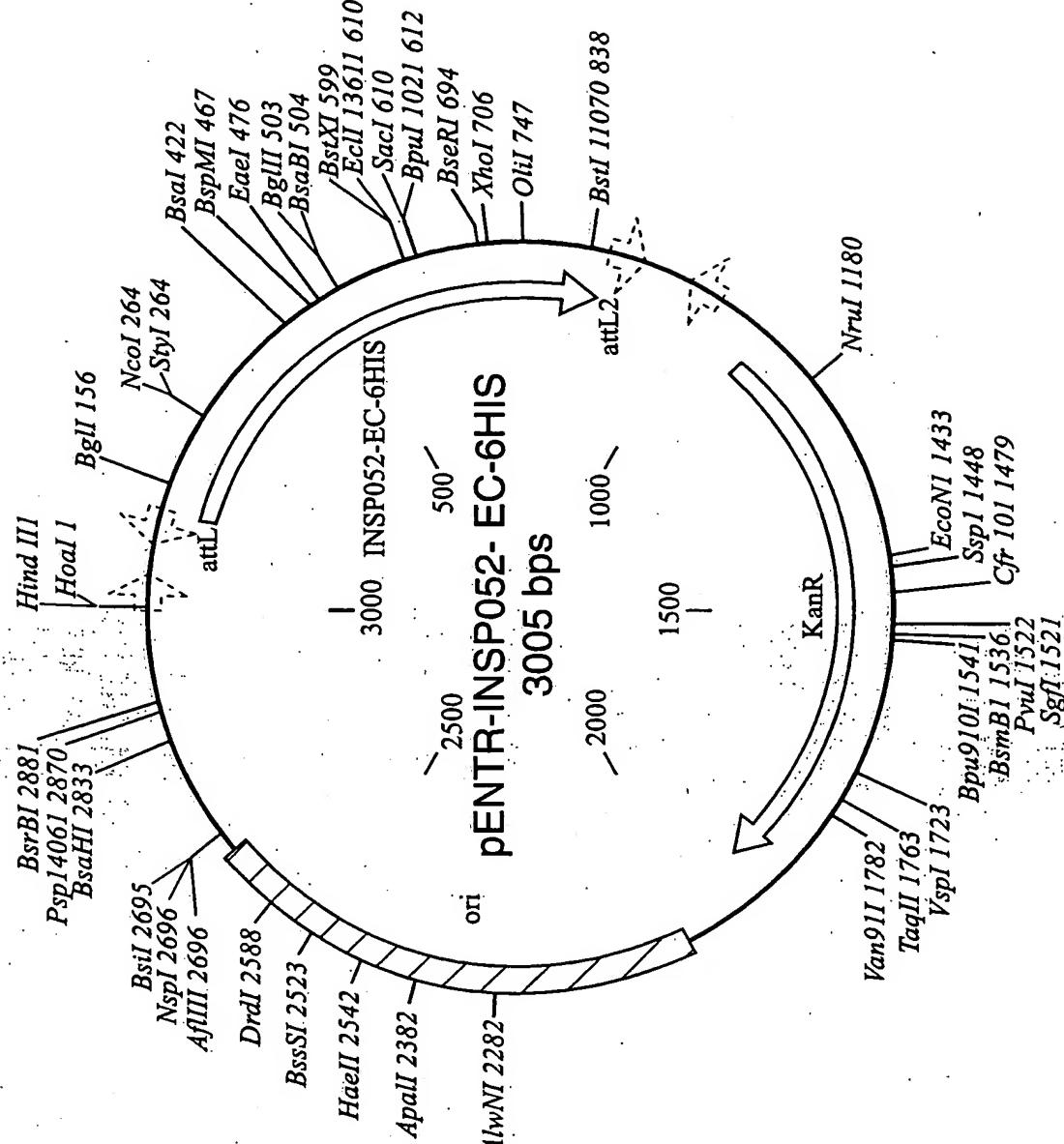


FIG. 8 (contd.)

## FIG. 9

Molecule: pEAK12d-INSP052-EC-6HIS, 7687 bps DNA Circular  
 File Name: pEAK12d-INSP052-6HIS.cm5, dated 21 Feb 2003

Description: Ligation of Cons-6His.SEQ into pEAK12d-attB1-attB2

Molecule Features:

Type	Start	End	Name	Description
REGION	2	595		pmb-ori
GENE	596	1519	AmpR	
REGION	1690	2795	EF-1a	
MARKER	2703			pEAK12F primer
REGION	2855	2887	attB1	
GENE	2888	3625	INSP052-EC-6HIS (aal-240)	
REGION	3629	3654	attB2	
MARKER	3656	C		pEAK12R primer
REGION	3661	4089		poly A/splice
GENE	4708	4090	C	PUROMYCIN resistance
REGION	4932	4709	C tK	tK promoter
REGION	5427	4933	C Ori P	
GENE	7479	5427	C EBNA-1	
REGION	7480	7679	sv40	

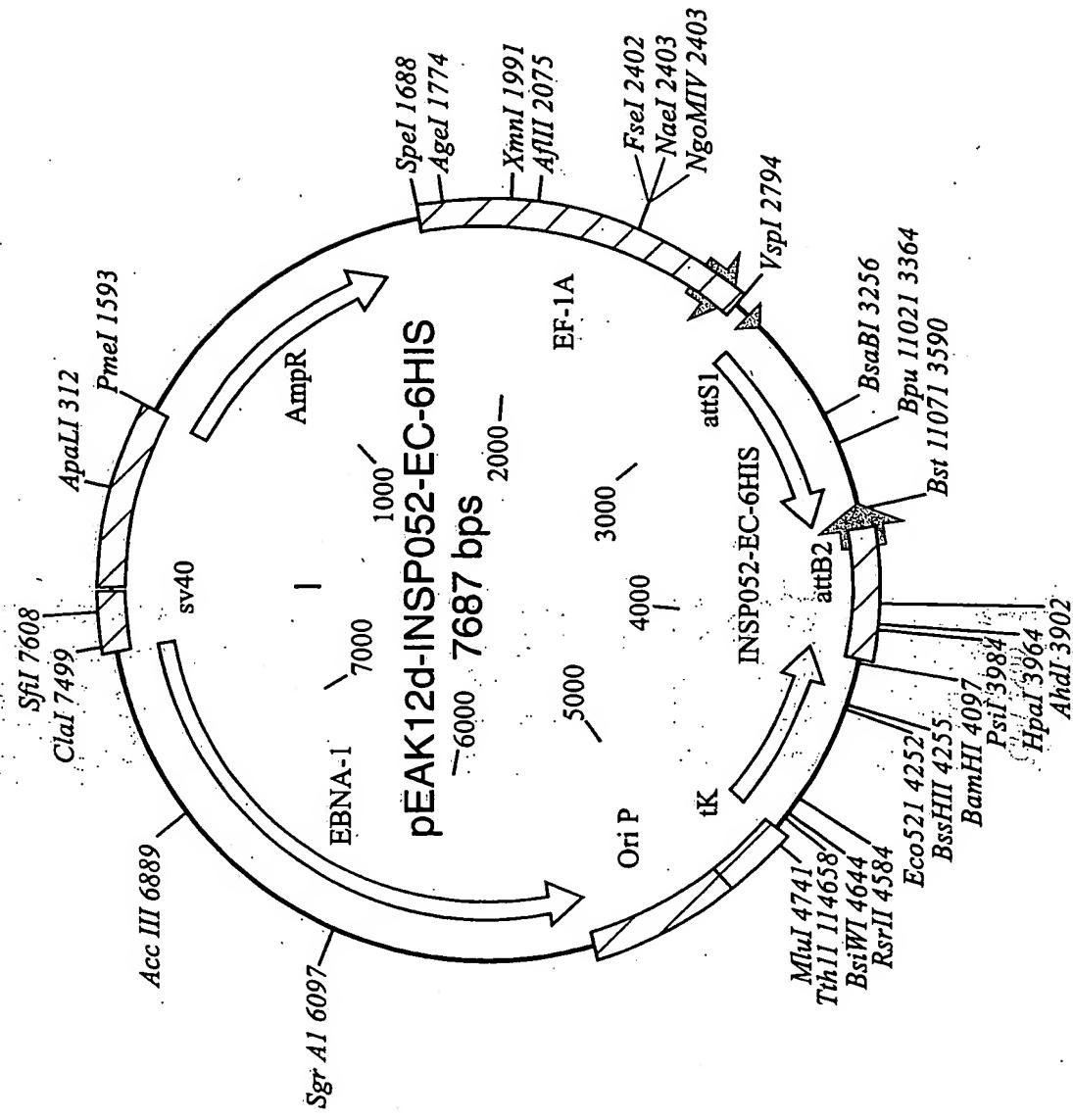


FIG. 9(contd.)

FIG. 10

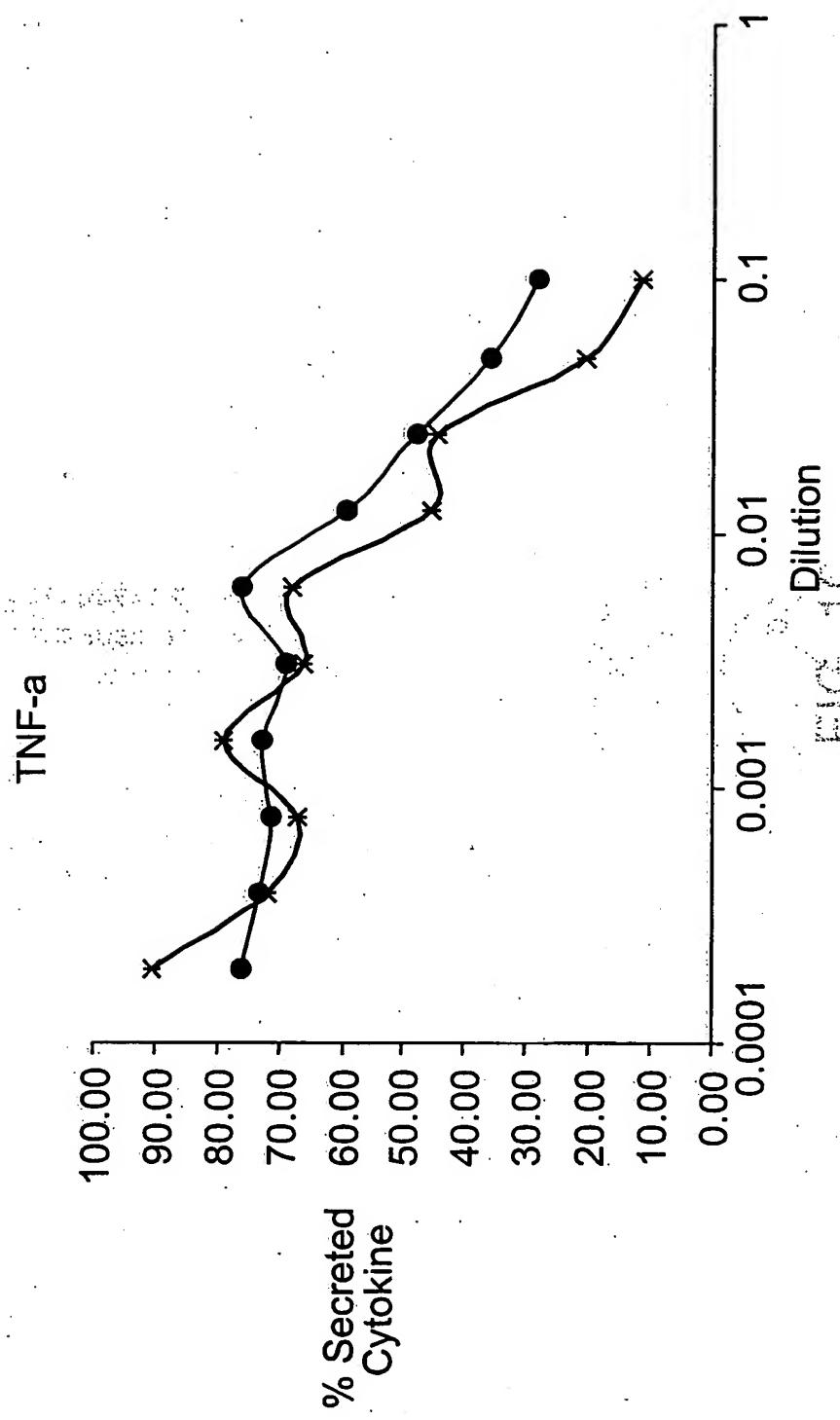


FIG. 11

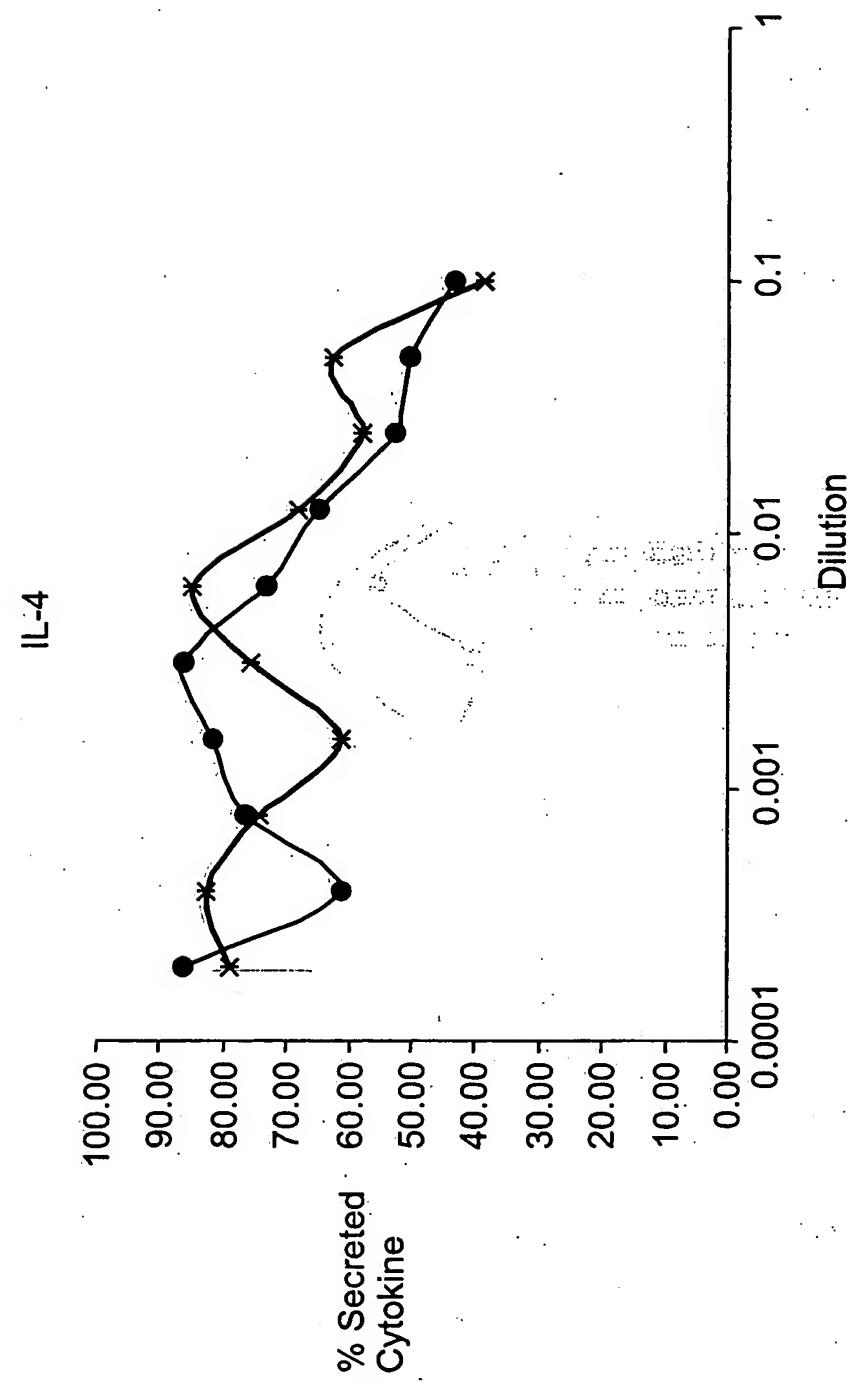
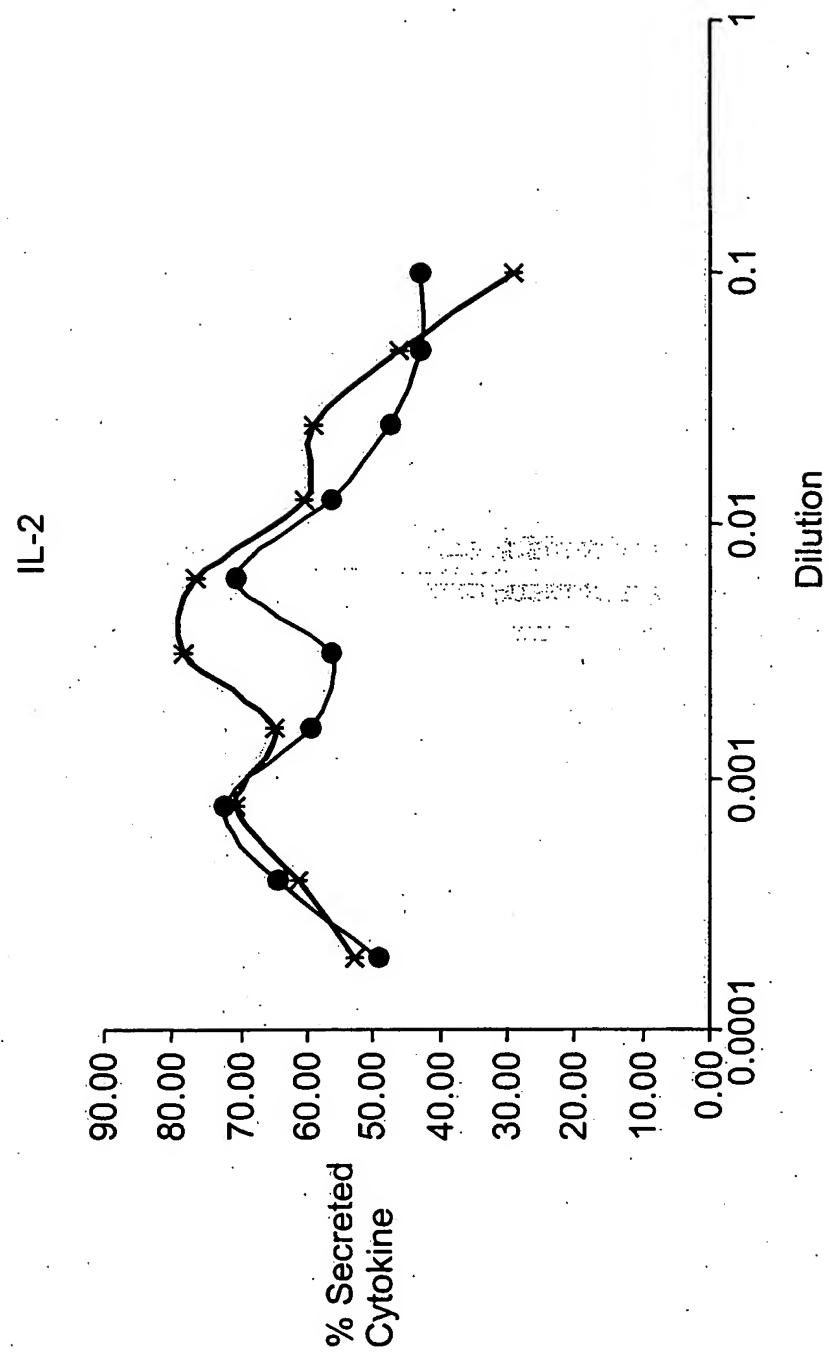


FIG. 12



19/22

FIG. 13A

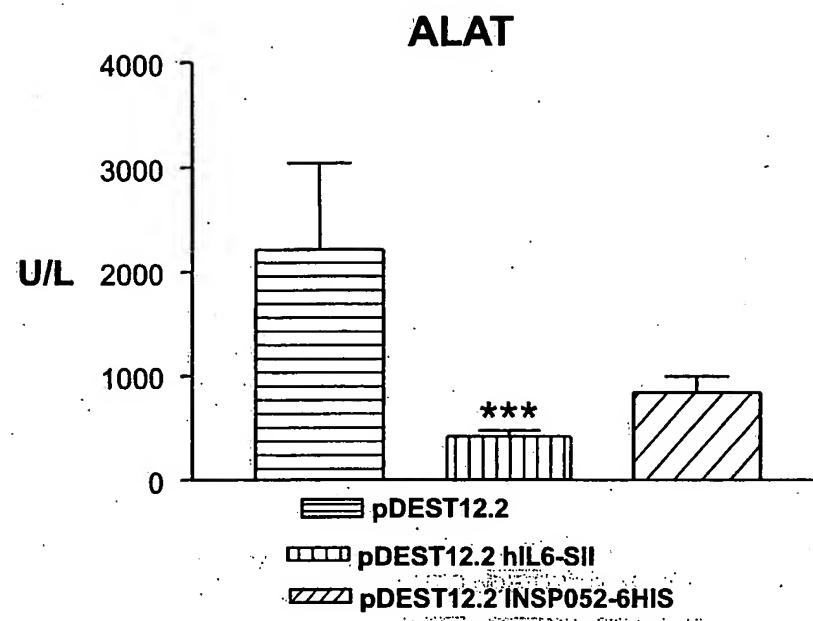
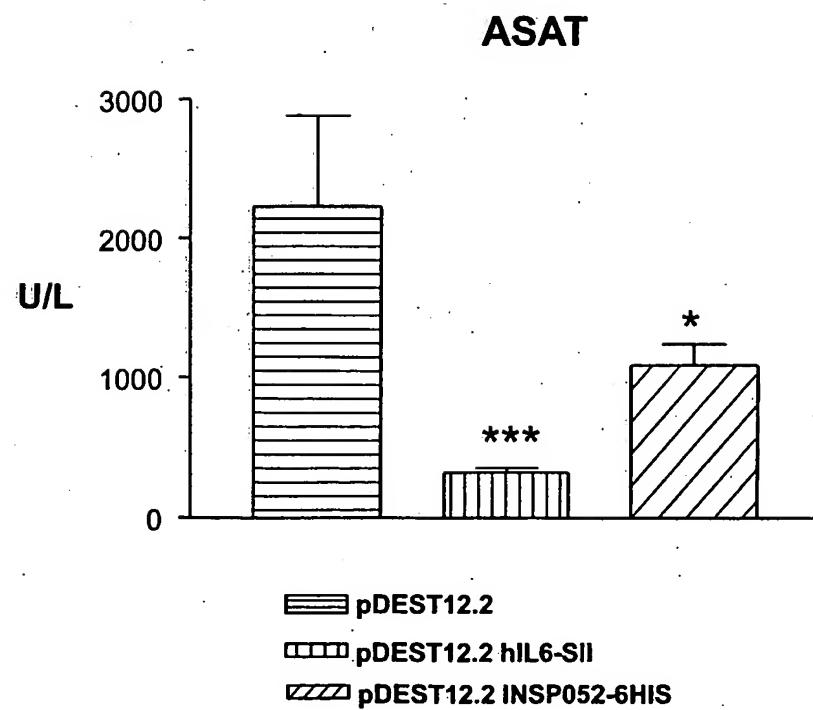


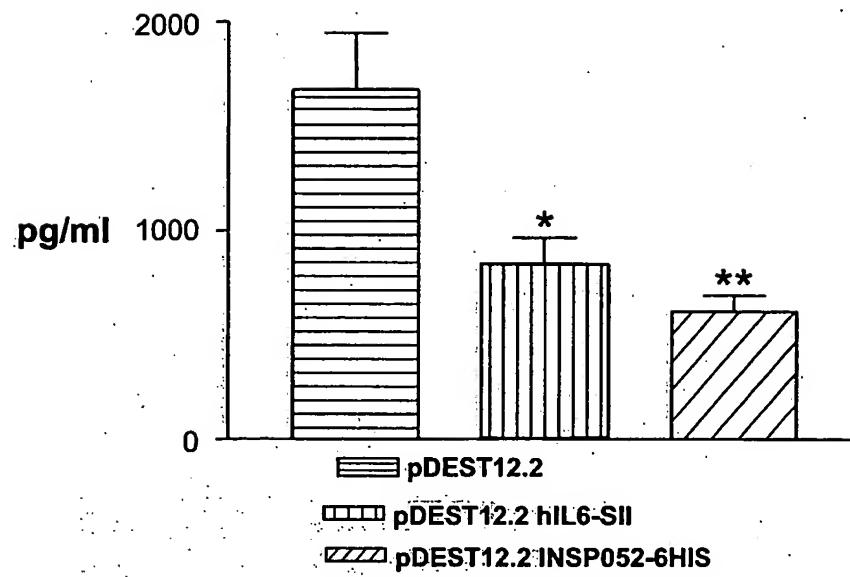
FIG. 13B



20/22

**FIG. 14A**

**TNF 1h30**



**FIG. 14B**

**mIL6 8h**

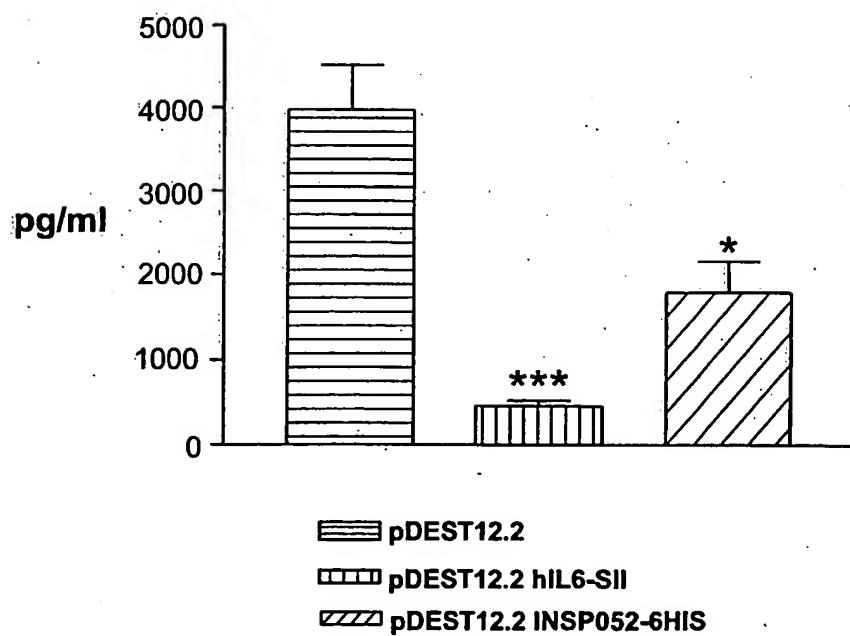


FIG. 15A

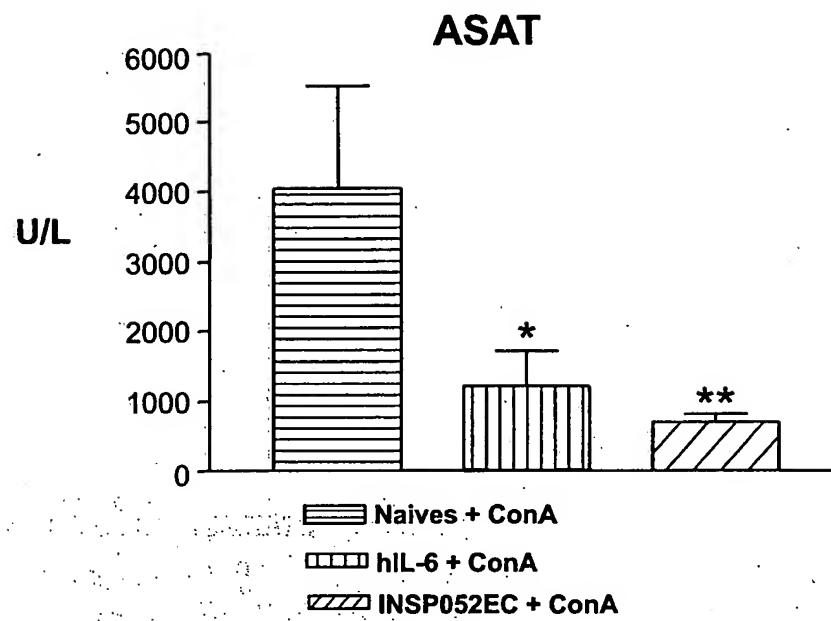
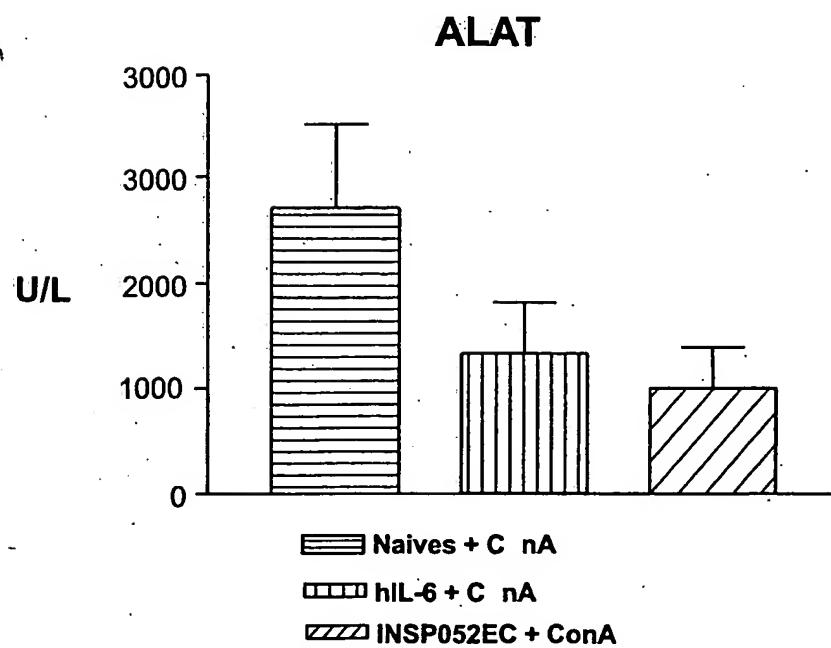


FIG. 15B



22/22

FIG. 15C

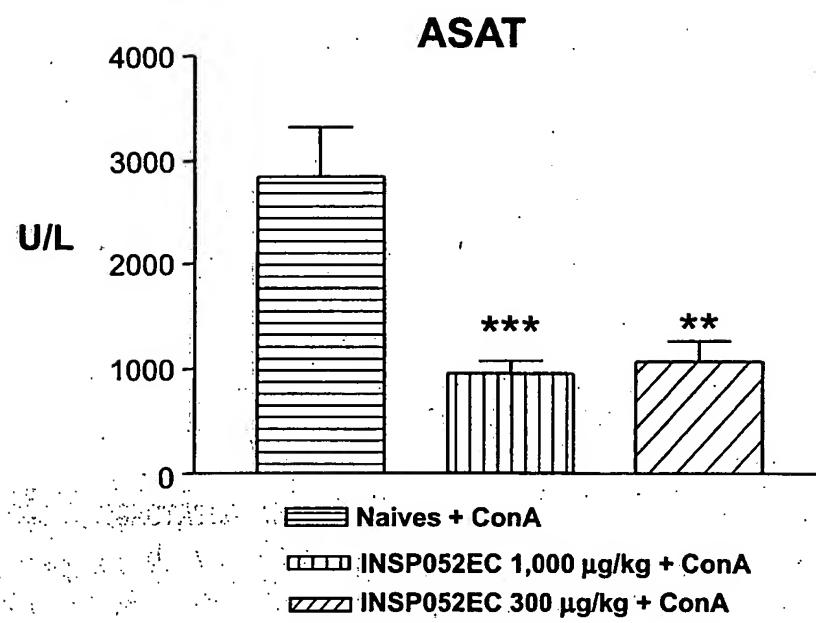


FIG. 15D

